



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

# The association of recombination events in the founding and emergence of subgenogroup evolutionary lineages of human enterovirus 71

### Citation for published version:

McWilliam Leitch, EC, Cabrerizo, M, Cardoso, J, Harvala, H, Ivanova, OE, Koike, S, Kroes, AC, Lukashev, A, Perera, D, Roivainen, M, Susi, P, Trallero, G, Evans, DJ & Simmonds, P 2012, 'The association of recombination events in the founding and emergence of subgenogroup evolutionary lineages of human enterovirus 71', *Journal of Virology*, vol. 86, no. 5, pp. 2676-2685. <https://doi.org/10.1128/JVI.06065-11>

### Digital Object Identifier (DOI):

[10.1128/JVI.06065-11](https://doi.org/10.1128/JVI.06065-11)

### Link:

[Link to publication record in Edinburgh Research Explorer](#)

### Document Version:

Publisher's PDF, also known as Version of record

### Published In:

Journal of Virology

### Publisher Rights Statement:

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

### General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# The Association of Recombination Events in the Founding and Emergence of Subgenogroup Evolutionary Lineages of Human Enterovirus 71

E. C. McWilliam Leitch,<sup>a</sup> M. Cabrerizo,<sup>b</sup> J. Cardoso,<sup>c</sup> H. Harvala,<sup>d</sup> O. E. Ivanova,<sup>e</sup> S. Koike,<sup>f</sup> A. C. M. Kroes,<sup>g</sup> A. Lukashev,<sup>e</sup> D. Perera,<sup>b</sup> M. Roivainen,<sup>h</sup> P. Susi,<sup>i</sup> G. Trallero,<sup>b</sup> D. J. Evans,<sup>j</sup> and P. Simmonds<sup>a</sup>

Centre for Infectious Diseases, University of Edinburgh, Summerhall, Edinburgh, United Kingdom<sup>a</sup>; Enterovirus Laboratory, National Centre for Microbiology, Carlos III Institute of Health, Majadahonda, Madrid, Spain<sup>b</sup>; Institute of Health and Community Medicine, University Sarawak Malaysia, Sarawak, Malaysia<sup>c</sup>; Specialist Virology Centre, Royal Infirmary of Edinburgh, Edinburgh, United Kingdom<sup>d</sup>; M.P. Chumakov Institute of Poliomyelitis and Viral Encephalitis, Moscow, Russia<sup>e</sup>; Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan<sup>f</sup>; Department of Medical Microbiology, Leiden University Medical Centre, Leiden, The Netherlands<sup>g</sup>; Intestinal Viruses Unit, National Institute for Health and Welfare, Helsinki, Finland<sup>h</sup>; Department of Virology, University of Turku, Turku, Finland<sup>i</sup>; and Department of Biological Sciences, University of Warwick, Coventry, United Kingdom<sup>j</sup>

**Enterovirus 71 (EV71) is responsible for frequent large-scale outbreaks of hand, foot, and mouth disease worldwide and represent a major etiological agent of severe, sometimes fatal neurological disease. EV71 variants have been classified into three genogroups (GgA, GgB, and GgC), and the latter two are further subdivided into subgenogroups B1 to B5 and C1 to C5. To investigate the dual roles of recombination and evolution in the epidemiology and transmission of EV71 worldwide, we performed a large-scale genetic analysis of isolates ( $n = 308$ ) collected from 19 countries worldwide over a 40-year period. A series of recombination events occurred over this period, which have been identified through incongruities in sequence grouping between the VP1 and 3Dpol regions. Eleven 3Dpol clades were identified, each specific to EV71 and associated with specific subgenogroups but interspersed phylogenetically with clades of coxsackievirus A16 and other EV species A serotypes. The likelihood of recombination increased with VP1 sequence divergence; mean half-lives for EV71 recombinant forms (RFs) of 6 and 9 years for GgB and GgC overlapped with those observed for the EV-B serotypes, echovirus 9 (E9), E30, and E11, respectively (1.3 to 9.8 years). Furthermore, within genogroups, sporadic recombination events occurred, such as the linkage of two B4 variants to RF-W instead of RF-A and of two C4 variants to RF-H. Intriguingly, recombination events occurred as a founding event of most subgenogroups immediately preceding their lineage expansion and global emergence. The possibility that recombination contributed to their subsequent spread through improved fitness requires further biological and immunological characterization.**

Enterovirus 71 (EV71) is one of the most frequently detected pathogenic human enteroviruses, responsible for large-scale epidemic occurrences of neurological disease throughout South-east Asia (51, 58). EV71 contains a single-stranded, positive-sense RNA genome and is classified as a member of species A (EV-A) in the *Enterovirus* genus of the *Picornaviridae* family (60). As with other EV-A enteroviruses, EV71 is transmitted by the fecal-oral route and normally causes subclinical or relatively mild, self-limiting infections, such as hand, foot, and mouth disease (HFMD) (65). However, unlike other members of EV-A, EV71 infections are associated in a small proportion of subjects with a wide array of severe disease presentations, including aseptic meningitis, encephalitis, and acute flaccid paralysis (AFP) (reviewed in references 29 and 58). Severe and fatal EV71 infections are predominantly found in young children, with male patients outnumbering female patients (12, 52). There has been a substantial increase in the frequency and severity of EV71 epidemics in recent years, particularly in the Asian Pacific region (5, 13, 58, 65), prompting urgent, ongoing investigations of the virological and host factors contributing to the apparently increasing pathogenicity of the virus (5, 11, 58).

Considerable insights into the evolution and molecular epidemiology of circulating strains and genotypes of EV71 from many of the most-affected countries have been obtained through analysis of structural genome regions, principally VP1. EV71 has been classified into a total of three genogroups (Ggs), designated GgA

to GgC (6), showing approximately 13 to 20% amino acid sequence divergence from each other in the VP1 region (6, 28) and estimated to have originated from a common ancestor as recently as 1941 (61). Several studies have investigated whether different genogroups or subgenogroups vary in their pathogenicity, which might then explain the variability in outcomes of EV71 infections in different decades and between continents. While whole-genome sequence comparisons of EV71 strains isolated from severe or fatal cases of EV71 infection showed no reproducible differences from those causing more mild infections (54, 57), specific associations of GgC2 variants with severe neurological disease and of B3 with HFMD or mild/inapparent infections were observed during an outbreak in Perth, Australia, in 1988, when both genogroups were cocirculating (35, 36). More recently, a greater likelihood of GgC5 to cause neurological complications than GgC4 was reported (44).

In the current study we genetically characterized a large num-

Received 19 August 2011 Accepted 21 December 2011

Published ahead of print 28 December 2011

Address correspondence to P. Simmonds, Peter.Simmonds@ed.ac.uk.

Supplemental material for this article may be found at <http://jvi.asm.org/>.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JVI.06065-11

TABLE 1 Sources and collection dates of survey specimens

Virus and country	Code	No. of isolates	Subgenogroup(s)	Yr(s) of isolation
<b>EV71</b>				
Croatia	HR	2	C4	2005
Finland	FI	16	C1, C2, C5	1994, 2000–2005, 2007–2009
Georgia	GE	1	C2	2007
Great Britain	GB	5	C2	2010
Iceland	IS	2	C1	2004
Japan	JP	62	B2, B4, B5, C1, C2, C4	1990, 1993, 1997–2010
Latvia	LV	1	C1	2003
Malaysia	MY	83	B3, B4, B5, C1	1997–1998, 2000, 2002, 2003, 2006, 2008
The Netherlands	NE	15	C1, C2	2007, 2010
Russia	RU	6	C1, C2	2000, 2001, 2007–2009
Spain	ES	7	C1, C2	1999, 2002, 2003, 2007
Total ( <i>n</i> = 11)		200		1990, 1993–1994, 1998–2010
<b>CVA16</b>				
Finland	FI	10		2000, 2001, 2003, 2005, 2006, 2008–2010
Great Britain	GB	1		2009
Iceland	IS	2		2002, 2004
Japan	JP	1		2004
Latvia	LV	1		2007
The Netherlands	NL	4		2008
Russia	RU	5		2008–2010
Slovak Republic	SK	2		2004–2005
Spain	ES	11		2000, 2003, 2006, 2008
Total ( <i>n</i> = 9)		37		2000–2010

ber of EV71 strains, collected from South and East Asia, Australia, and Europe, in the VP1 region and a distal region of the genome (part of the 3D polymerase-encoding region [3Dpol]) for identification of recombination events. Recombination is a frequently documented phenomenon in picornaviruses and contributes to their evolutionary diversity as well as a means to acquire new phenotypic traits from acquisition of novel combination of structural and nonstructural genes and 5'-untranslated region (5'-UTR) sequences.

Recombination in picornaviruses was first observed between serotypes of poliovirus in vaccine recipients (8, 20, 39) and more recently in a wide range of human enteroviruses (1, 15, 16, 21, 31, 41, 43, 45, 48, 56), foot-and-mouth disease virus (FMDV), and teschoviruses (55) and, more recently, parechoviruses (3, 4). In each virus, recombination breakpoints concentrate in the 2A region, although further sites occur in P2 and P3 genes and in the 5'-UTR (1, 3, 16, 30, 33, 49, 67). In contrast, phylogenies of the more divergent capsid-encoding genes VP1, VP2, and VP3 are congruent with each other (and correspond to their serotypic classification) (31, 42, 49), leading to the idea that sequence diversification in structural gene regions is largely uncoupled from that of nonstructural genes (31, 32, 37, 55). The contribution of recombination to the evolution and molecular epidemiology of EV71 and its relationship to diversification of capsid-encoding genes is the focus of the current study.

The study used a large data set of newly acquired and published sequences for the VP1 (capsid) and 3Dpol (nonstructural) gene regions of EV71 and coxsackievirus A16 (CVA16). The identification of recombination events was assisted through the assignment

of 3Dpol sequences of EV71 isolates to individual recombinant forms, as developed for the analysis of EV species B viruses (37, 38). The study provides convincing evidence for recombination events in the founding of most subgenogroup lineages over an evolutionary time scale comparable to that observed within EV-B serotypes. These analyses further our understanding of the molecular epidemiology of EV71 and its varied clinical manifestations.

## MATERIALS AND METHODS

**Samples.** A total of 193 isolates from 11 countries collected between the years 1990 and 2010 (Table 1) were obtained from internationally distributed referral centers. The following convention was used to name isolates: two-letter country code and isolate number/two-letter city or region abbreviation/3Dpol clade/year of collection (e.g., MY40/Sw/A/06 for isolate number 40 referred from Sarawak in Malaysia, isolated in 2006, and belonging to the 3Dpol clade A [defined below]). Sequences obtained in the current study were supplemented with 5 previously unpublished sequences from Russia collected between 2000 and 2009 and with 110 published sequences of complete genomes from 10 countries that were collected between 1970 and 2009, including the EV71 prototype strain BrCr-CA-70. All sequences in the current study were obtained from laboratory-passaged virus stocks, as were the published complete genome sequences of EV71 and CVA16 incorporated in the analysis. This greatly reduced the likelihood of sampling mixed virus populations that might, for example, have originated from coinfections of the study subjects with different EV71 or CVA16 strains.

**Amplification of VP1 and 3Dpol regions and nucleotide sequencing.** RNA extraction and nested reverse transcription-PCRs (RT-PCRs) were performed as previously described but using newly designed enterovirus primers specific for EV71 and CVA16. These amplified a 1,055-bp region of the VP1 gene and a 759-bp region of the 3Dpol gene. Primers com-

prised the following: for VP1, OS (outer sense; position 2268 in the poliovirus type 3 sequence; K01392), CCN TGG ATH AGY AAC ACN CAY T; OAS (outer antisense; position 3604), GAR AAR CTR ACY GGR TAG TGY TTT CT; IS (inner sense; position 2332), TNA SNA TYT GGT AYC ARA CAN AYT; and IAS (inner antisense; position 3409), ACR TAD ATD GCN CCN GAY TGY TG. For 3Dpol, the primers included OS (position 5830), GSA CYA TGA TGT AYA AYT TYC CHA C; OAS (position 7045), GGN GTC ATD GTY ARN CCR TAY TGY TT; IS (position 6261), ATG AGY ATR GAR GAN GCN TGY TAY G; and IAS (position 7195), TCY TTN GTC CAD CGR ATR GAY TCR T. Amplicons were directly sequenced using BigDye (Applied Biosystems), and the inner sense or antisense primer and nucleotide sequences were aligned with the Simmonic sequence package, version 1.9 (56; <http://www.virus-evolution.org/Software>). Confirmatory repetitions of RT-PCRs and sequencing were performed for all isolates that showed incongruences in tree positions between the VP1 and 3Dpol regions, as described below.

**Phylogenetic analysis.** Bootstrapped maximum likelihood trees for VP1 and 3Dpol regions were generated using RAxML with the GTRGAMMA model (general time-reversible plus gamma distribution for rates over sites, with all model parameters estimated using RAxML) and 100 bootstraps (59). Regression analysis and investigation of geographical and temporal aspects of recombination were performed using maximum composite likelihood (MCL) distances (calculated by using MEGA) between sequences.

A Bayesian Markov chain Monte Carlo (MCMC) method implemented in the BEAST package, version 1.53 (18), was used to estimate temporal phylogenies and rates of evolution (19). Individual data sets prepared for BEAST analysis (corresponding to genogroups, subgenogroups, or 3Dpol clades used to define RFs) were checked for recombination using the program Genetic Algorithm Recombination Detection (GARD) or the single breakpoint (SBP) method for larger VP1 data sets (GgC and RF-W) in the Datamonkey package, which provides an interface to the HyPhy program (27, 47). Further testing was performed on each data set by using several algorithms implemented in the RDP package (RDP, GENCONV, MaxChi, Chimaera, SiScan, and 3Seq [34]).

For BEAST analysis we used constant population size as a prior, while for selected data sets other tree priors were used (exponential growth, Bayesian Skyline) to determine the effect priors on analysis outcomes. Two independent runs for each set were analyzed using the SRD06 model of substitution (53), with chain lengths of 50 million or 100 million and a relaxed molecular clock model that allows evolutionary rates to vary among lineages. All other parameters were optimized during the burn-in period. Output from BEAST was analyzed using the program TRACER (<http://beast.bio.ed.ac.uk/Tracer>), and the results of each duplicate were compared. RF succession dynamics were examined using a maximum clade credibility (MCC) tree visualized in FigTree following annotation in Tree-Annotator.

**Nucleotide sequence accession numbers.** All newly generated sequences obtained in this study were submitted to GenBank and were assigned the accession numbers [HQ676156](#) to [HQ676288](#) (VP1) and [HQ676289](#) to [HQ676487](#) (3Dpol).

## RESULTS

**Phylogeny of EV71 VP1 and 3Dpol genome regions.** A total of 198 EV71 isolates from 11 countries and 37 CVA16 isolates (Table 1) were analyzed in the current study concurrently with the prototype sequence BrCr-CA-70 (U22521; designated US01/CA/Q/70 in this study) and 109 previously published full-genome sequences of EV71 and 8 of CVA16 (see Table S1 in the supplemental material). EV71 and CVA16 sequences were amplified and sequenced in the VP1 region between positions 2458 and 3345 (numbering based on BrCr-CA-70). All EV71 isolates analyzed in the current study clustered with other EV71 sequences to form a monophyletic group separate from other EV-A serotypes/

types (data not shown). The isolates assembled into the previously designated genogroups B2 to B5, C1, C2, C4, and C5; those from Europe were exclusively GgC1, GgC2, GgC3, and GgC5 (see Fig. S1 in the supplemental material).

For recombination analysis, sequences from the 3' end of the genome (within the 3Dpol coding region) were determined for each of the 198 isolates analyzed in VP1, and the data set was combined with available complete genome sequences of EV71 and CVA16 (see Fig. S1 in the supplemental material). Sequences in the 3Dpol region formed a series of bootstrap-supported clades; EV71 3Dpol clades comprised groups A, D, E, G, H, L, Q, T, V, W, and Y, while clades C, I, J, M, O, and S were identified among CVA16 sequences. Individual genogroups were associated with specific 3Dpol clades in the majority of cases.

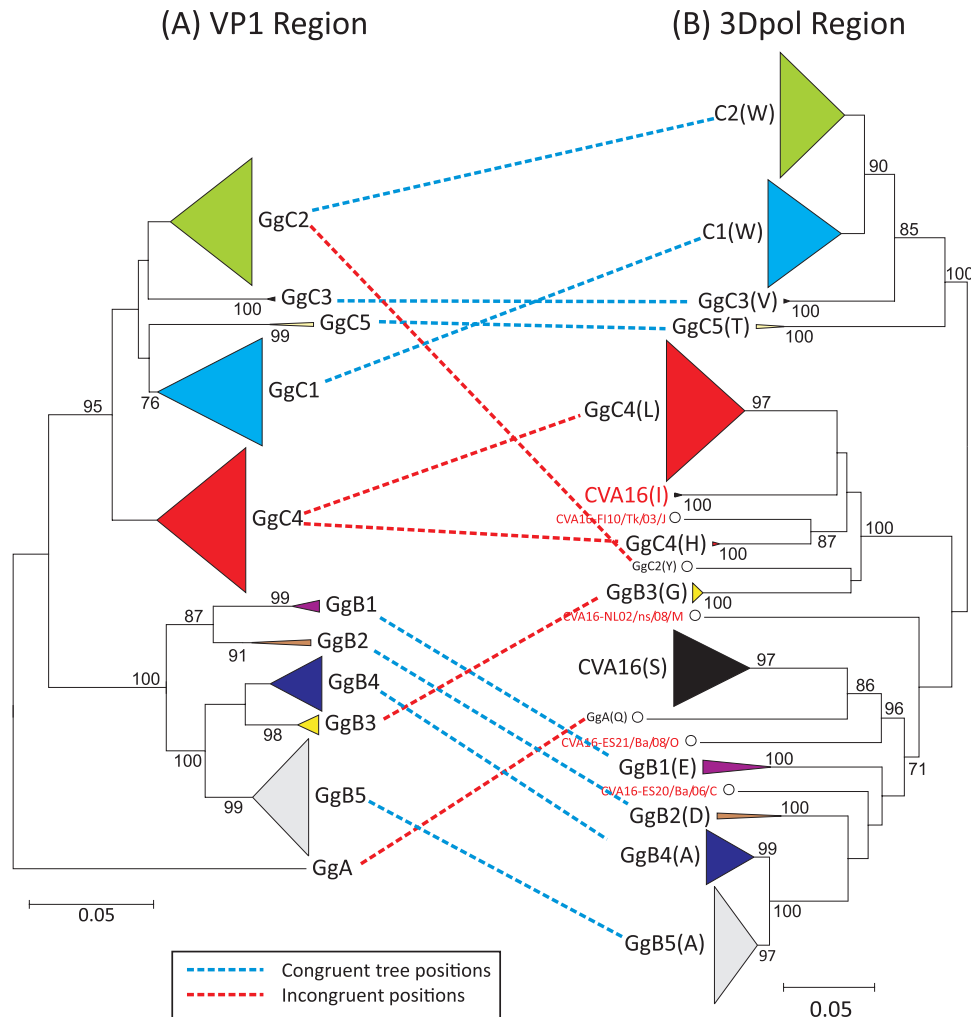
Supporting these phylogenetic assignments, the pairwise distributions of sequence distances between variants in both VP1 and 3Dpol fell into a series of discontinuous ranges (see Fig. S2A in the supplemental material). For example, the minimum value in the VP1 distribution separating the second and third distributions of approximately 19% corresponds to the threshold value separating intra- from intergenogroup distances between GgA, GbB, and GgC (6) (see Fig. S2A). The lower threshold value (10 to 11%) corresponded to the previously designated subgenogroup boundaries (6). An analogous division of 3Dpol sequences into the 15 clades by phylogenetic analysis was similarly supported by their distribution of pairwise distances (see Fig. S2B). The low point in the distribution, 0.13, corresponded closely to the threshold value dividing distances between and within the phylogenetically defined clades.

3Dpol classification provides the means to identify and quantify recombination events within data sets through identification of incongruences between phylogenetically supported clades in different genome regions, as used in previous analyses of EV-B serotypes (37, 38). 3Dpol gene sequences of the 15 RFs of EV71 and CVA16 indeed imperfectly mapped onto the phylogeny of the VP1 region and revealed several phylogeny violations indicative of recombination (Fig. 1). By branch rotation to maximize tree branching orders and by ignoring any phylogenetic grouping without 70% or more bootstrap support, trees from the two regions revealed both branching order differences within EV71 (red dotted lines) as well as highly interspersed groupings of CVA16 sequences in the 3Dpol tree (red sequence labels) that contrasted with the consistent outgroup position of CVA16 in the VP1 region (Fig. 1).

As specific examples, GgB3 grouped with B4 and B5 in VP1 but adopted a distinct tree position in 3Dpol closest to C4 and CVA16 variants. The branching order of GgA was similarly incongruent. Within genogroups, while the majority of sequences were associated with a specific 3Dpol group, such as C2 with W and C4 with L, individual variants within each showed phylogenetically distinct 3Dpol sequences similarly indicative of recombination. Each of the CVA16 groups identified as separate RFs originated through at least 6 further recombination events in their evolutionary histories. This analysis depicts the minimum number of recombination events required to resolve the phylogenies of the available sequences. The interspersed nature of other EV-A 3Dpol sequences in the phylogenetic tree (data not shown) entails a much larger number of further recombination events over the longer period of EV species A evolution.

**Sequence diversification in VP1 and 3Dpol regions.** EV71





**FIG 1** Phylogeny of VP1 and 3Dpol regions of EV71 for study subjects and those previously determined (listed in Table 1; see also Table S1 in the supplemental material). Clades were identified by bootstrap analysis (values of  $\geq 70\%$ ) by maximum likelihood analysis as implemented using RAxML. The sizes of the triangles are proportional to the number of sequences within each clade (bootstrap resampling values are shown on branches). By branch rotation to maximize visual congruence of the two trees and the use of bootstrap values of  $\geq 70\%$  to define phylogenetic groupings, the minimum numbers of incongruent phylogeny relationships (depicted by red dotted lines) were determined (clades showing congruent branching orders labeled with blue dotted lines). The 3Dpol region includes available sequences from CVA16; these are labeled red, as their interspersed positions in 3Dpol are invariably incongruent with their outgroup position in VP1 (not included in the tree on the left).

VP1 variability was predominantly restricted to synonymous sites indicative of predominantly neutral sequence change. Using the data available on the collection dates of samples in a Bayesian MCMC method (BEAST [19]) with a relaxed molecular clock for substitution rate calculations, substitution rates, and times of most recent common ancestor (TMRCA) of different genogroups and 3Dpol clades defining the largest RFs were calculated. Each data set was analyzed for internal recombination prior to BEAST analysis. GARD/SBP detected no recombination events in any data set, while detection of recombinant sequences using RDP-implemented methods was infrequent and inconsistent between methods (see Table S4 in the supplemental material). To investigate whether these possibly recombinant sequences influenced the validity of the MCMC analysis, one of the sequences in VP1 identified as recombinant by three methods (TW13) was excluded from the GgC data set, and CN01 was removed from the clade L 3Dpol sequence data set. Substitution rates and TMRCA

were compared with the full data sets (see Table S5 in the supplemental material). Effectively identical rates and dates were calculated for each comparison, indicating that intraclade recombination had no effect on the validity of the BEAST analysis.

The crude evolutionary rate of evolution for the whole data set of VP1 sequences was  $7.2 \times 10^{-3}$  substitutions/site/year (high-probability distribution [HPD] range,  $6.2 \times 10^{-3}$  to  $8.3 \times 10^{-3}$ ) (Table 2), with a predicted TMRCA of all extant EV71 lineages of 80 years (HPD, 58 to 106 years). Analysis of individual genogroups within EV71, GgB, and GgC revealed comparable evolutionary rates of  $6.2 \times 10^{-3}$  substitutions/site/year (HPD range,  $5.2 \times 10^{-3}$  to  $7.3 \times 10^{-3}$ ) for GgB and  $5.0 \times 10^{-3}$  substitutions/site/year (HPD range,  $4.3 \times 10^{-3}$  to  $5.8 \times 10^{-3}$ ) for GgC and TMRCA of 36 years (HPD, 35 to 38 years) and 34 years (HPD, 25 to 37 years), respectively. Similar substitution rates and more recent TMRCA (8 to 32 years) were observed among subgenogroups (Table 2).

TABLE 2 Rates of sequence change and TMRCA by MCMC analysis

Genogroup and geographic set	Gg or RF	<i>n</i> <sup>c</sup>	Divergence <sup>a</sup>				Regression (R)		MCMC (BEAST) <sup>b</sup>			
			Nucleotide		aa				Substitution rate (10 <sup>−3</sup> ) <sup>d</sup>		TMRCA <sup>e</sup>	
			VP1	3Dpol	VP1	3Dpol	VP1	3Dpol	VP1	3Dpol	VP1	3Dpol
Whole data set												
All	All	308	0.18	—	0.02	—	0.14	—	7.2 (6.2–8.3)	—	79.9 (58.2–105.5)	—
Europe	All	58	0.09	—	0.01	—	0.58	—	3.5 (1.9–5.1)	—	36.2 (19.4–58.1)	—
Asia	All	244	0.17	—	0.02	—	0.32	—	6.4 (5.2–7.8)	—	61.4 (45.8–79)	—
Individual genogroups												
All	GgB	108	0.07	—	0.01	—	0.94	—	6.2 (5.2–7.3)	—	35.9 (35–37.5)	—
All	GgC	199	0.13	—	0.02	—	0.77	—	5.0 (4.3–5.8)	—	30.7 (24.8–37.6)	—
Asia	GgC	137	0.11	—	0.01	—	0.77	—	5.4 (4.5–6.2)	—	27.4 (23.6–31.7)	—
Individual subgenogroups												
All	GgB4	28	0.03	—	0.01	—	0.7	—	6.8 (2.2–13)	—	8.3 (5.4–12.4)	—
All	GgB5	60	0.02	—	0.00	—	0.88	—	7.3 (4.6–10.4)	—	9.1 (8–10.6)	—
All	GgC1	55	0.05	—	0.01	—	0.86	—	3.2 (2.1–4.3)	—	31.2 (22.9–41.1)	—
All	GgC2	65	0.05	—	0.01	—	0.75	—	4.7 (3.1–6.2)	—	21.0 (15.4–28.8)	—
All	GgC4	74	0.04	—	0.01	—	0.75	—	6.0 (4.6–7.4)	—	15.4 (12.8–18.4)	—
Individual RF groups												
All	RF-A	88	0.05	0.06	0.01	0.01	0.92	0.95	6.4 (4.8–8.1)	6.7 (5.3–8.3)	13.3 (11.7–15.1)	13.9 (12.3–15.6)
All	RF-W	119	0.10	0.09	0.01	0.02	0.35	0.76	5.1 (3.8–6.5)	4.3 (2.7–5.9)	30.5 (23.3–38.9)	34.9 (24.5–47.9)
Asia	RF-W	61	0.10	0.09	0.01	0.02	0.64	0.75	5.2 (4.1–6.3)	5.7 (4.2–7.5)	26.1 (22.5–29.7)	24.5 (20.4–28.3)
Europe	RF-W	55	0.08	0.08	0.01	0.01	0.59	0.57	3.7 (2.2–5.4)	1.1 (0.4–1.9)	30.4 (17–47.3)	85.1 (28.8–161.3)
All	RF-L	72	0.04	0.04	0.01	0.01	0.85	0.89	5.9 (4.6–7.2)	5.4 (4.0–6.8)	15.5 (12.9–18.6)	18.8 (13.5–25.3)

<sup>a</sup> Mean pairwise *P* distances. —, values were not calculated for the 3Dpol region for groups in which 3Dpol sequences were not monophyletic. aa, amino acids.

<sup>b</sup> The mean value is based on two independent analyses; substitution rates and TMRCA were not calculated for groups where 3Dpol sequences were not monophyletic.

<sup>c</sup> The number of sequences analyzed in each set.

<sup>d</sup> Frequency of substitutions per site per year (with HPD range shown in parentheses).

<sup>e</sup> Time before the present of the most recent common ancestor (in years).

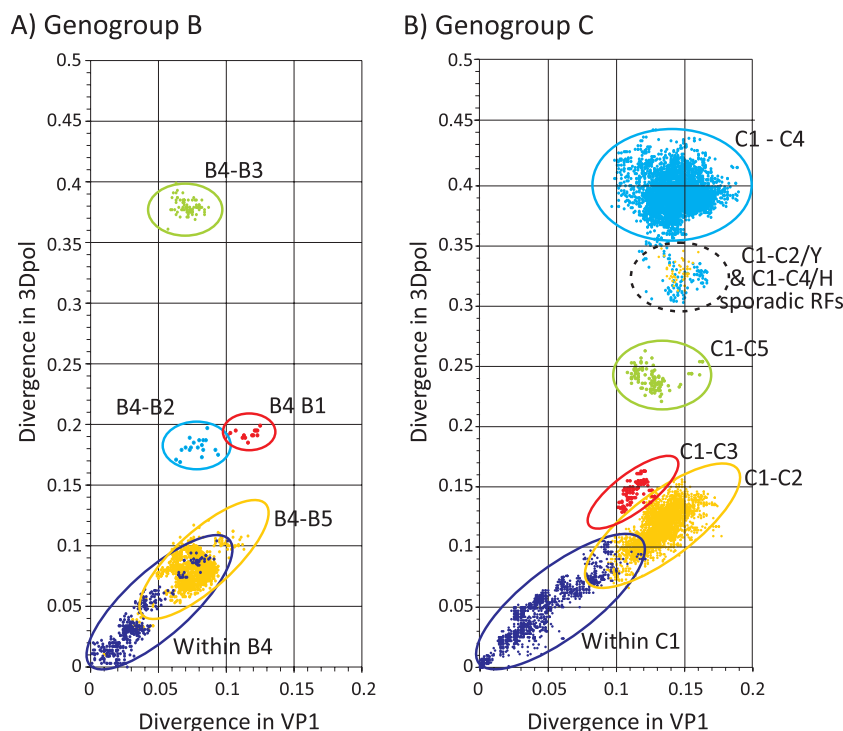
For the three larger RF groups (RF-A, RF-L, and RF-W), remarkably similar substitution rates for individual clades were determined in the 3Dpol region (Table 2). For example, the substitution rate for RF-A was  $6.4 \times 10^{-3}$  substitutions/site/year (HPD range,  $4.8 \times 10^{-3}$  to  $8.1 \times 10^{-3}$ ) in the VP1 region and  $6.7 \times 10^{-3}$  substitutions/site/year (HPD range,  $5.3 \times 10^{-3}$  to  $8.3 \times 10^{-3}$ ) in the 3Dpol region. Furthermore, both genome regions analyzed provided consistent estimates for the TMRCA of each individual RF group analyzed; using again the example of RF-A, TMRCA of 13.3 years (HPD, 11.7 to 15.1) and 13.9 years (HPD range, 12.3 to 15.6) for VP1 and 3Dpol, respectively, were observed. These data provided robust estimates for the dates of the recombination events that created each RF.

To determine whether the prior on the tree in the BEAST analysis influenced estimates of substitution rate or TMRCA, analyses were repeated using exponential growth and Bayesian Skyline analyses for the larger data sets (GgB and GgC in VP1 and RF-L and RF-W in 3Dpol) (see Table S3 in the supplemental material). Each prior produced extremely similar values for substitution rate and TMRCA in each data set. Sequence relationships between dated samples collected over a wide temporal range primarily informed these calculations rather than tree prior.

The similar dynamics of sequence drift between VP1 and 3Dpol regions and the disruptive effect of recombination were apparent through comparison of pairwise distances in the two genome regions of the better-represented subgenogroups (Fig. 2; examples of B4 and C1 are shown). Discontinuities in the distributions

of pairwise distances supported the hypothesis of widespread recombination during EV71 evolution. Within-genogroup GgB4 and GgC1 comparisons showed an approximate straight line correlation between VP1 and 3Dpol divergence (dark blue points), with a positive gradient of approximately 1 and a *y* intercept approximating 0. This was consistent with the similar substitution rates observed between the two regions by Bayesian MCMC analysis (Table 2). Both showed evidence for continuous and discontinuous sequence distributions compared with other subgenogroups. Distributions of pairwise distances between B4 and B5 and between C1 and C2 overlapped, and their similar trajectories to within-subgenogroup (B4 and C1) comparisons demonstrated a process of ongoing divergence without recombination. B4 and B5 indeed retained the same RF group, A, as did C1 and C2 (with RF-W). Comparison of B4 and C1 sequences to other subgenogroups revealed quite distinct relationships. For example, C2 and C4 were similarly divergent from C1 in VP1 (comparing the distributions of yellow and light blue points), but 3Dpol sequences from C4 were substantially divergent (pairwise distance of around 0.4), an observation that can only be accounted for by a recombination event, in this case to RF group L.

Equivalent recombination events are implied by the other intersubgenogroup comparisons comprising B4 (A) to B1 (E), B2 (D), and B3 (G) and C1 (W) to C4 (L), C5 (T), and likely C3 (V) (RF designations in parentheses). Comparison of C1 to C2 and C4 further revealed the separate placement of RF-Y and RF-H, away from the main groups, confirming the occurrence of further spo-



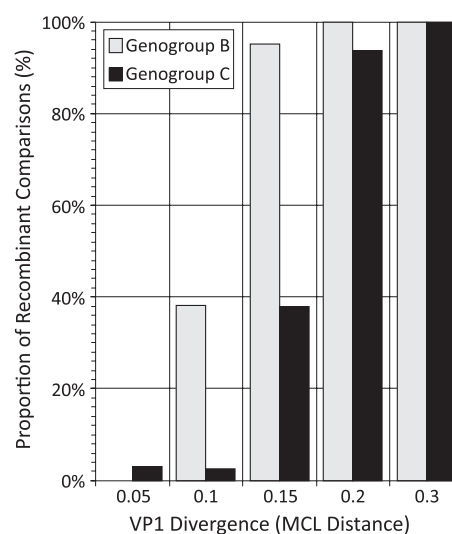
**FIG 2** Relationships between sequence divergence (MCL pairwise distances) in the VP1 region (x axis) and in 3Dpol (y axis) among (A) genogroup B (A) and genogroup C (B) EV71 sequences. In the examples shown, sets of pairwise distances in both regions between GgB4 and C1 to other variants within GgB and GgC are depicted. Note that subgenogroups C2 and C4 both contained within them single sporadic recombinants; these account for the additional groupings of data points encircled by the dotted line.

radic recombination events. Indeed, pairwise comparison of each subgenogroup to each within gB and gC subgroups (30 comparisons [data not shown]) confirmed that only B4/B5 and C1/C2 were consistently equidistant in the two genome regions; these were additionally the only two pairs of subgenogroups that shared the same 3Dpol group (A and W, respectively).

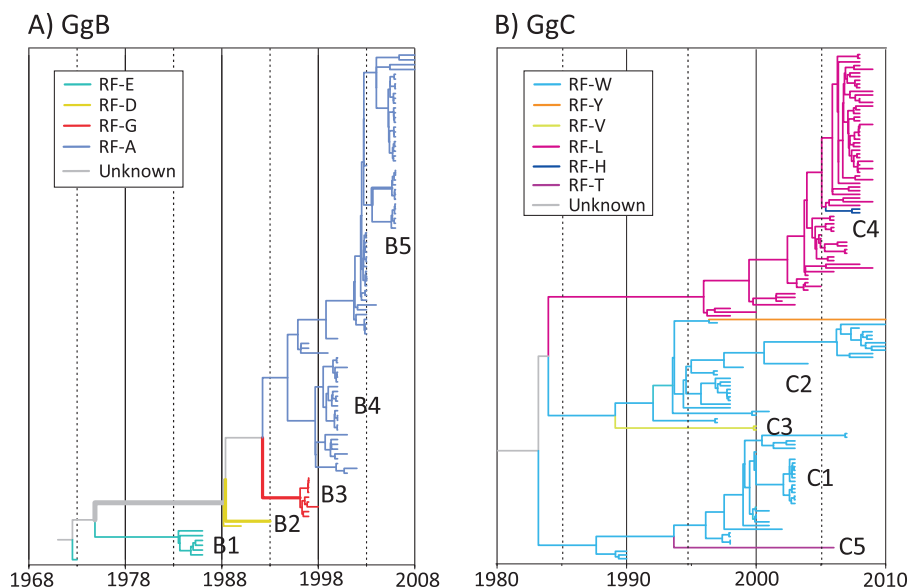
**Temporal correlates of recombination in GgB and GgC.** To estimate the relationship between virus diversification and the occurrence of recombination in EV71, sequences of each isolate within EV species B and C were compared in the VP1 region to estimate evolutionary divergence and in the 3Dpol region to identify shared or different 3Dpol groupings (Fig. 3). Pairs of isolates with different 3Dpol groups were considered to have undergone recombination. The proportion of isolate comparisons with different 3Dpol groups increased with increasing VP1 divergence in both genogroups B and C (Fig. 3). As previously performed for EV-B serotypes (38), interpolation of the 50% value, combined with the substitution rate for VP1 (Table 2 [ $7.2 \times 10^{-3}$  substitutions/site/year]) enabled estimates of approximate half-lives of genogroups B and C EV71 to be made. Values of 0.085 and 0.135 predicted half-lives of 5.9 and 9.4 years for GgB and -C, respectively. This difference is consistent with the observation of lower divergence between recombinant sequences in GgB in the analysis of VP1 and 3Dpol sequences (Fig. 3) and the prolonged existence of subgenogroups C1 and C2, which retained a long-term association with the 3Dpol clade W, despite their substantial sequence divergence in VP1 (pairwise distances ranging from approximately 0.10 to 0.17 [Fig. 2]).

The occurrence of recombination was mapped onto MCMC-

generated time-correlated trees (Fig. 4) to estimate when individual recombination events occurred. GgB and GgC displayed different temporal dynamics and patterns of RF succession. For GgB (Fig. 4A), a stepwise, time-related correlation of recombination with VP1 divergence was observed, encompassing isolates collected over a 35-year period. As identified previously, each VP1



**FIG 3** Association between VP1 sequence divergence (the maximum value is shown for each bar on the x axis) and the proportion of recombinant comparisons (i.e., belonging to different 3Dpol clades) for GgB and GgC.



**FIG 4** MCMC tree of the VP1 sequences of GgB (A) and GgC (B) from the Asian Pacific region visualized using FigTree and plotted on a temporal y-axis scale using their sampling dates. Branches are color coded (see the key in each panel) according to the recombination group of individual sequences and their reconstructed ancestors.

subgenogroup corresponded to a single RF group, with the exception of GgB4 and GgB5 (both RF-A). The consecutive replacements of B1 by B2, B2 by B3, and B3 by B4 each involved viruses bearing distinct 3Dpol region sequences (RF-E to RF-D to RF-G to RF-A). The date of recombination events could be estimated as lying between the MRCA shared by the original RF and the new RF and the isolation date of the first clinical sample with the new RF. Thus, it can be estimated that recombination of the most recent RF group, RF-A, occurred between 1992 and 1997. The subsequent emergence of the associated subgenogroups, GgB4 and GgB5, was characterized by large-scale, short-term outbreaks occurring with a periodicity of approximately 3 years. However, unlike earlier outbreaks (1973 to 1998) involving RF groups E, D, and G, the 3-year spacing of these outbreaks was not immediately preceded by a recombination event even when VP1 divergence was extensive enough to designate a separate subgenogroup (GgB4 and GgB5).

GgC comprised three major lineages that diverged from a common ancestor in 1983 (HPD, 1978 to 1986 [Fig. 4B]). Two of the VP1 lineages in GgC contained one or two sporadic monophyletic RF groups interspersed within one major RF group. It is likely that these sporadic RF groups originated from a single (datable) recombination event, and each were minor and short-lived components of the circulating virus population. The lineage containing GgC1 (RF-W) persisted for at least 17 years (1990 to 2007), and GgC5 (RF-T) emerged as a sporadic RF group in 2006, having undergone a recombination event sometime between 1994 and 2006. The first clinical sample of the second lineage (mainly GgC2, RF-W) was collected in 1997, and this lineage has persisted for 13 years, up to the end of the study period in 2010. This lineage also contained two GgC3 samples belonging to RF-V that were isolated in 2000 and likely recombined between 1989 and 2000. A further sporadic RF group (Y) that appeared in 2010 probably recombined with its unknown second parental strain between 1996 and 2010, exhibiting a long quiescent period before emergence in

2010. The third lineage was comprised entirely of GgC4 sequences, which persisted for 11 years (between 1998 and 2009), displayed time-correlated divergence, and consisted predominantly of RF-L with the exception of two isolates of the sporadic group RF-H (isolated in 2008); this group probably recombined between 2005 and 2008, emerging in 2008 after a short quiescent period. RF-H did not appear to replace the parental RF-L group up to the end of the study in 2010.

## DISCUSSION

**Detection of recombination in EV71.** This study applied methods developed in previous genetic analyses of EV-B viruses to reexamine the occurrence and dynamics of recombination in EV71. EV71 is considered to be the most pathogenic of the currently circulating enteroviruses worldwide, with infections associated with outbreaks of HFMD and serious neurological disease in Southeast Asia. Understanding the underlying interactions between virus evolution and population susceptibility and the biological basis for its severe disease associations are major research priorities, as is the development of preventative or treatment strategies for its control.

In the current study, detection of recombination events was achieved through identification of bootstrap-supported clades by phylogenetic analysis of the 3Dpol region. These groupings were used to categorize EV71 variants into a series of RFs whose assignments were supported by parallel analyses of pairwise distances (see Fig. S2 in the supplemental material); variants within the same RF group showed pairwise distances in the 3Dpol region of  $<0.19$  and of  $>0.19$  between those in different groups. By these criteria, the EV71 3Dpol sequences in this study assembled into 11 clades, and 5 of these clades comprised over 96% of the total non-structural sequences analyzed. Therefore, the repertoire of 3Dpol RFs was much more restricted in EV71 than that found in species B EVs (37, 38), where 119 E11 variants collected over a 14-year period could be assigned to 43 RF groups, 89 E9 samples into 23



RF groups, and 240 European E30 samples into 26 RF groups. The underlying reasons for this difference in diversity between viruses in EV species A and B are unclear.

Using this classification, comparison of phylogenetic trees from VP1 and 3Dpol regions showed a large number of phylogenetic incongruities indicative of recombination (Fig. 1). Most striking was the interspersed position of 3Dpol sequences from CVA16 variants included in the analysis. As these by definition take an outlier position in the VP1 region, each occurrence therefore represents a recombination event in the evolutionary history of either EV71 or CVA16. Also striking was the difference of tree position of GgB3 and GgC4 in VP1 and in 3Dpol; both grouped away from other members of the same genogroup (10, 17, 24, 66). However, the assertion that these variants show evidence for intertypic recombination, such as GgC4 with CVA16, is not supported by the analysis in the current study. GgC4 variants are, with one exception, assigned to RF-H, a 3Dpol grouping that is distinct from its closest neighbor, RF-L, assigned to CVA16. Although occupying neighboring positions in the tree, their assignment to different RF groups therefore provides evidence against the frequently proposed specific recombination event between GgC4 and CVA16. Subgenogroup B3 showed an analogous change in tree position, as previously described (10), although again without evidence for intertypic recombination, since its 3Dol group (G) is also not shared with any other EV71 or CVA16 variants (Fig. 1) or other EV-A serotype (data not shown).

This tree comparison also revealed the existence of two sporadic recombinants, variants of EV71 with 3Dpol assignments different from the rest of the subgenogroup within which they are classified. These comprise the C2 variant characterized in the current study, JP17/Ac/Y/10, assigned as RF-Y instead of the majority, RF-W (Fig. 1), and the C4 variants CN23/Sz/H/08 and CN19/Bj/H/08, assigned as RF-H instead of RF-L (Fig. 1). The other evident phylogeny violation is the outlier position of the single GgA sequence in VP1 and its inlier position in 3Dpol.

Indeed, rather than representing defined recombination events between serotypes, or indeed recombination between lineages within EV71, the discordant sequence relationships between genome regions in EV71 closely resemble a quite different process of largely independent evolution of genome regions that has been previously observed in species B enteroviruses (32, 37, 38). For echoviruses 9, 11, and 30, evolutionary lineages identified within the capsid region show evidence for recombination with a larger pool of nonstructural region variants that show no systematically closer genetic relationship with one serotype than with any other. In the case of E9, E11, and E30, the 92 RF groups were thus fully interspersed with each other and with 3Dpol groups of other EV-B serotypes, with only very rare occurrences of shared RF groups between different serotypes or indeed lineages within a serotype (37, 38). The observed scattered positions of CVA16 and EV71 3Dpol-associated clades and the lack of shared RF assignments of different subgenogroups observed in the current study are indeed precise mirrors of the pattern observed in EV-B.

Distributions of pairwise distances in the two genome regions (Fig. 2) provided evidence for further recombination events in the evolution of EV71. This analysis method demonstrated abrupt discontinuities in distributions of pairwise distances between the two genome regions, indicative of recombination events accompanying the founding of most subgenogroups within GgB and GgC. This conclusion is consistent with independent evolutionary

pathways of structural and nonstructural genome regions identified in EV-B (32). Indeed, only pairwise comparisons between members of the same subgenogroup, or between B1/B2 and C4/C5, showed equality in divergence (i.e., the gradient of the set of pairwise distances approximated unity), which was entailed in their similar substitution rates (Table 2). All other comparisons between subgenogroups revealed disproportionately high and nonlinear greater divergence in the 3Dpol region, indicative of recombination. These plots of pairwise distances help visualize the differing degrees of overall divergence in different genome regions described previously (64) and which are so evident on divergence scans and boot-scanning (10, 17, 24, 66).

**Time scale of recombination events.** By measuring the relationship between VP1 divergence and occurrence of recombination, as previously carried out for EV-B serotypes (38), it was possible to calculate approximate half-lives for individual RFs of EV71 (Fig. 3). Estimates of 5.9 and 9.4 years for Gg B and GgC overlapped the range observed previously in the EV-B serotypes E9, E30, and E11 (with half-lives of 1.3, 3.1, and 9.8 years, respectively [38]). However, the major EV71 RFs (A, E, G, L, and W) showed decades-long circulation without recombination, a pattern also observed for the E11 RF-DU, which continued to circulate throughout a study period from 1996 to 2008 (38). Variability in recombination frequency is likely shaped by differing viral epidemiologies that govern opportunities for coinfection and generation of hybrid viruses to occur. They may also be influenced by different compatibility restrictions, which would dictate the likelihood of replication-competent viruses being generated by recombination. The marked differences in recombination frequencies observed between human parechovirus (HPeV) type 1 (4 years) and HPeV type 3 (20 years) (7) may indeed be a manifestation of constraints limiting compatibility between viruses with different cell entry mechanisms and, potentially, cellular tropisms (22).

More-precise identification of the individual recombination events in EV71 was achieved through the use of time-correlated trees and superimposition of branching points in the VP1 tree that most parsimoniously accounted for the RF designations in descendant sequences (Fig. 4). These phylogenetic reconstructions for genogroups B and C additionally place recombination events into the differing evolutionary trajectories of GgB and GgC that have been characterized previously (61, 63). GgB is characterized by a series of successive emergence and extinctions of the B1 through B5 subgenogroups over the period from 1970 to the present day, with outbreaks occurring throughout South and East Asia in a cycle of approximately every 3 years (1993, 1997, 2000, 2003, 2006, and 2008) (9, 23, 46, 58, 62). The association between the founding of each lineage and recombination is clearly evident from the phylogenetic analysis. This provides the means to estimate dates within fairly narrow windows for their occurrence. These tree-based estimates are consistent with the TMRCA estimates from diversity/substitution rate calculations in both the VP1 and 3Dpol regions (Table 2) (61).

In contrast, it has been established that different subgenogroups of GgC EV71 variants have coexisted for at least 25 years, with three separate lineages (C1, C2, and C5) emerging in the 1980s and continuing to circulate to the present (5, 58, 61). This contrasting pattern is exemplified in a previous characterization of isolates of EV71 in Malaysia; GgB variants were isolated only during periodic major outbreaks, whereas GgC isolates were de-

tected sporadically, both during and between epidemics (46). Whether or by what mechanism these contrasting epidemiologies contribute to the differing recombination frequencies of GgB and GgC (Fig. 3) remains to be determined.

**Evolution of EV71.** It is well established that the longer-term genetic and antigenic diversification of human enteroviruses and the existence of serologically distinct EV types are key factors in their evolutionary success and ongoing ubiquitous presence in human populations. What is rather less clear is the underlying mechanisms and selection pressures involved in the generation of new serotypes, as this process has to date not been directly observed. In the case of EV71, it has been hypothesized that the evolution of capsid genes, particularly VP1, represents an equivalent, immunologically driven process of diversification (61, 63) and that the successive appearance of novel genogroups and subgenogroups is favored by an absence of preexisting immunity to them. The genetically diverse variants of EV71 may thus be precursors in their eventual further diversification into new serotypes in the future. This hypothesis is, however, not clearly supported by existing genetic and antigenic comparisons of EV71 genogroups (14, 24–26, 40, 61, 63). First, sequence divergence between genogroups and subgenogroups occurs overwhelmingly at silent sites, indicative of purifying or neutral evolutionary drift (14, 24, 61). Screening of large data sets of VP1 sequences provides little if any evidence for any sites in the capsid coding region being subjected to the positive selection that would typically be observed in sites under strong immunological pressure. Consistent with these analyses, there is little evidence from serological cross-neutralization experiments for genogroup- or subgenogroup-specific antibodies, nor indeed the existence of measurable antigenic diversity between EV71 isolates (24, 26, 28, 40, 63). However, infections with GgB variants induce higher levels of neutralizing antibodies than GgC, supporting an idea that the emergence of GgC may have been assisted by its intrinsically lower immunogenicity than GgB (40).

The alternative hypothesis is that the emergence and turnover of EV71 genogroups simply represent random processes of emergence and extinction of lineages without underlying natural selection, as proposed for other enteroviruses (2, 37, 50). Random fixation of individual variants with no fitness advantage within a population is indeed more likely when population sizes are small. Such conditions may occur during the periodic bottlenecks in a population size that inevitably occur in viruses that cause acute infections and show epidemic cycles of transmission. These instances of very rapid turnover and complete population replacements have been extensively documented, with a similar lack of evidence for antigenic replacement. However, as observed for EV71, such turnover is frequently associated with recombination events occurring during the founding of new evolutionary lineages. Understanding whether the partial or complete replacement of nonstructural gene regions associated with such recombination events provides a replicative or immunological selective advantage to the virus and thus drives the diversification of EV71 and other enteroviruses is a key unanswered question. Future biological and immunological investigation of both capsid region sequence change and recombination are clearly required if we are really to understand the evolution of enteroviruses and indeed of other nonenveloped RNA viruses.

## ACKNOWLEDGMENTS

We thank staff at the following universities and hospitals for technical assistance with virus isolation: Department of Virology, University of Turku, Turku, Finland; Intestinal Viruses Unit, National Institute for Health and Welfare, Helsinki, Finland; Gurutze Rubio, Cruces Hospital, Bilbao, Spain; Manuel Omeñaca, Miguel Servet Hospital, Zaragoza, Spain; Nuria Rabella, Santa Cruz y San Pablo Hospital, Barcelona, Spain; Carmen Perez, Dr. Negrin Hospital, Las Palmas de Gran Canaria, Spain; T. P. Eremeeva, M.P. Chumakov Institute of Poliomyelitis and Viral Encephalitis, Moscow, Russia (the latter supported in part by the Polio Eradication Initiative through the European Office of the World Health Organization); Seiya Yamayoshi (Neurovirology Project, Tokyo Metropolitan Institute of Medical Science, Kamikitazawa, Setagaya-ku, Tokyo, Japan). We additionally thank Setsuko Izuka (Shimane Prefectural Institute of Public Health and Environmental Science, Nishihamasado, Matsue, Shimane, Japan), Teruo Yamashita and Hiroko Minagawa (Laboratory of Virology, Department of Microbiology and Medical Zoology, Aichi Prefectural Institute of Public Health, Nagare, Tsujimachi, Kita-ku, Nagoya, Aichi, Japan), Katsumi Mizuta (Department of Microbiology, Yamagata Prefectural Institute of Public Health, Tokamachi, Yamagata, Japan), and Hidekazu Nishimura (Virus Research Center, Sendai Medical Center, Miyagino, Miyagino-ku, Sendai, Miyagi, Japan) for providing samples.

We are very grateful to Sam Lycett and Andrew Rambaut (Institute of Evolutionary Biology, University of Edinburgh) for valuable discussions and assistance with phylogenetic and BEAST analyses.

This study was funded by a project grant from the Wellcome Trust.

## REFERENCES

- Andersson P, Edman K, Lindberg AM. 2002. Molecular analysis of the echovirus 18 prototype: evidence of interserotypic recombination with echovirus 9. *Virus Res.* 85:71–83.
- Bailey JL, et al. 2009. Phylogeography of circulating populations of human echovirus 30 over 50 years: nucleotide polymorphism and signature of purifying selection in the VP1 capsid protein gene. *Infect. Genet. Evol.* 9:699–708.
- Benschop K, et al. 2010. Comprehensive full length sequence analyses of human parechoviruses; diversity and recombination. *J. Gen. Virol.* 91:151–154.
- Benschop KS, Williams CH, Wolthers KC, Stanway G, Simmonds P. 2008. Widespread recombination within human parechoviruses: analysis of temporal dynamics and constraints. *J. Gen. Virol.* 89:1030–1035.
- Bible JM, Pantelidis P, Chan PKS, Tong CYW. 2007. Genetic evolution of enterovirus 71: epidemiological and pathological implications. *Rev. Med. Virol.* 17:371–379.
- Brown BA, Oberste MS, Alexander JP, Jr., Kennett ML, Pallansch MA. 1999. Molecular epidemiology and evolution of enterovirus 71 strains isolated from 1970 to 1998. *J. Virol.* 73:9969–9975.
- Calvert J, et al. 2010. The recombination dynamics of human parechoviruses; investigation of type-specific differences in frequency and epidemiological correlates. *J. Gen. Virol.* 91:1229–1238.
- Cammack N, Phillips A, Dunn G, Patel V, Minor PD. 1988. Intertypic genomic rearrangements of poliovirus strains in vaccinees. *Virology* 167:507–514.
- Cardosa MJ, et al. 2003. Molecular epidemiology of human enterovirus 71 strains and recent outbreaks in the Asia-Pacific region: comparative analysis of the VP1 and VP4 genes. *Emerg. Infect. Dis.* 9:461–468.
- Chan YF, Abubakar S. 2004. Recombinant human enterovirus 71 in hand, foot and mouth disease patients. *Emerg. Infect. Dis.* 10:1468–1470.
- Chang LY, et al. 2008. HLA-A33 is associated with susceptibility to enterovirus 71 infection. *Pediatrics* 122:1271–1276.
- Chen KT, Chang HL, Wang ST, Cheng YT, Yang JY. 2007. Epidemiologic features of hand-foot-mouth disease and herpangina caused by enterovirus 71 in Taiwan, 1998–2005. *Pediatrics* 120:E244–E252.
- Chen TC, et al. 2006. Combining multiplex reverse transcription-PCR and a diagnostic microarray to detect and differentiate enterovirus 71 and coxsackievirus A16. *J. Clin. Microbiol.* 44:2212–2219.
- Chen X, et al. 2010. Analysis of recombination and natural selection in human enterovirus 71. *Virology* 398:251–261.

15. Chevaliez S, et al. 2004. Molecular comparison of echovirus 11 strains circulating in Europe during an epidemic of multisystem hemorrhagic disease of infants indicates that evolution generally occurs by recombination. *Virology* 325:56–70.
16. Cuervo NS, et al. 2001. Genomic features of intertypic recombinant Sabin poliovirus strains excreted by primary vaccinees. *J. Virol.* 75:5740–5751.
17. Ding NZ, et al. 2009. Appearance of mosaic enterovirus 71 in the 2008 outbreak of China. *Virus Res.* 145:157–161.
18. Drummond AJ, Ho SY, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4:e88.
19. Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214.
20. Furione M, et al. 1993. Polioviruses with natural recombinant genomes isolated from vaccine-associated paralytic poliomyelitis. *Virology* 196: 199–208.
21. Guillot S, et al. 2000. Natural genetic exchanges between vaccine and wild poliovirus strains in humans. *J. Virol.* 74:8434–8443.
22. Harvala H, Wolthers K, Simmonds P. 2010. Parechoviruses in children: understanding a new infection. *Curr. Opin. Infect. Dis.* 23:224–230.
23. Hosoya M, et al. 2006. Genetic diversity of enterovirus 71 associated with hand, foot and mouth disease epidemics in Japan from 1983 to 2003. *Pediatr. Infect. Dis. J.* 25:691–694.
24. Huang SC, et al. 2008. Appearance of intratypic recombination of enterovirus 71 in Taiwan from 2002 to 2005. *Virus Res.* 131:250–259.
25. Huang SW, et al. 2009. Reemergence of enterovirus 71 in 2008 in Taiwan: dynamics of genetic and antigenic evolution from 1998 to 2008. *J. Clin. Microbiol.* 47:3653–3662.
26. Huang YP, et al. 2010. Genetic diversity and C2-like subgenogroup strains of enterovirus 71, Taiwan, 2008. *Virol. J.* 7:277.
27. Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD. 2006. Automated phylogenetic detection of recombination using a genetic algorithm. *Mol. Biol. Evol.* 23:1891–1901.
28. Kung S-H, et al. 2007. Genetic and antigenic analyses of enterovirus 71 isolates in Taiwan during 1998–2005. *Clin. Microbiol. Infect. Dis.* 13:782–787.
29. Lee TC, et al. 2009. Diseases caused by enterovirus 71 infection. *Pediatr. Infect. Dis. J.* 28:904–910.
30. Lindberg AM, Andersson P, Savolainen C, Mulders MN, Hovi T. 2003. Evolution of the genome of human enterovirus B: incongruence between phylogenies of the VP1 and 3CD regions indicates frequent recombination within the species. *J. Gen. Virol.* 84:1223–1235.
31. Lukashev AN. 2005. Role of recombination in evolution of enteroviruses. *Rev. Med. Virol.* 15:157–167.
32. Lukashev AN, et al. 2005. Recombination in circulating human enterovirus B: independent evolution of structural and non-structural genome regions. *J. Gen. Virol.* 86:3281–3290.
33. Lukashev AN, Lashkevich VA, Koroleva GA, Ilonen J, Hinkkanen AE. 2004. Recombination in uveitis-causing enterovirus strains. *J. Gen. Virol.* 85:463–470.
34. Martin DP, et al. 2010. RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* 26:2462–2463.
35. McMinn P, et al. 2001. Phylogenetic analysis of enterovirus 71 strains isolated during linked epidemics in Malaysia, Singapore, and Western Australia. *J. Virol.* 75:7732–7738.
36. McMinn P, Stratov I, Nagarajan L, Davis S. 2001. Neurological manifestations of enterovirus 71 infection in children during an outbreak of hand, foot, and mouth disease in western Australia. *Clin. Infect. Dis.* 32: 236–242.
37. McWilliam Leitch EC, et al. 2009. Transmission networks and population turnover of echovirus 30. *J. Virol.* 83:2109–2118.
38. McWilliam Leitch EC, et al. 2010. Evolutionary dynamics and temporal/geographical correlates of recombination in the human enteroviruses, echovirus 9, 11, and 30. *J. Virol.* 84:9292–9300.
39. Minor PD, John A, Ferguson M, Icenogle JP. 1986. Antigenic and molecular evolution of the vaccine strain of type 3 poliovirus during the period of excretion by a primary vaccinee. *J. Gen. Virol.* 67:693–706.
40. Mizuta K, et al. 2009. Cross-antigenicity among EV71 strains from different genogroups isolated in Yamagata, Japan, between 1990 and 2007. *Vaccine* 27:3153–3158.
41. Norder H, Bjerregaard L, Magnius LO. 2002. Open reading frame sequence of an Asian enterovirus 73 strain reveals that the prototype from California is recombinant. *J. Gen. Virol.* 83:1721–1728.
42. Oberste MS, Maher K, Pallansch MA. 2004. Evidence for frequent recombination within species human enterovirus B based on complete genomic sequences of all thirty-seven serotypes. *J. Virol.* 78:855–867.
43. Oberste MS, Penaranda S, Maher K, Pallansch MA. 2004. Complete genome sequences of all members of the species human enterovirus A. *J. Gen. Virol.* 85:1597–1607.
44. Ooi MH, et al. 2007. Human enterovirus 71 disease in Sarawak, Malaysia: a prospective clinical, virological, and molecular epidemiological study. *Clin. Infect. Dis.* 44:646–656.
45. Oprisan G, et al. 2002. Natural genetic recombination between co-circulating heterotypic enteroviruses. *J. Gen. Virol.* 83:2193–2200.
46. Podin Y, et al. 2006. Sentinel surveillance for human enterovirus 71 in Sarawak, Malaysia: lessons from the first 7 years. *BMC Public Health* 6:180.
47. Pond SL, Frost SD, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21:676–679.
48. Santti J, Harvala H, Kinnunen L, Hyypia T. 2000. Molecular epidemiology and evolution of coxsackievirus A9. *J. Gen. Virol.* 81:1361–1372.
49. Santti J, Hyypia T, Kinnunen L, Salminen M. 1999. Evidence of recombination among enteroviruses. *J. Virol.* 73:8741–8749.
50. Savolainen C, Hovi T, Mulders MN. 2001. Molecular epidemiology of echovirus 30 in Europe: succession of dominant sublineages within a single major genotype. *Arch. Virol.* 146:521–537.
51. Schmidt NJ, Lennette EH, Ho HH. 1974. An apparently new enterovirus isolated from patients with disease of the central nervous system. *J. Infect. Dis.* 129:304–309.
52. Schuffenecker I, et al. 2011. Epidemiology of human enterovirus 71 infections in France, 2000–2009. *J. Clin. Virol.* 50:50–56.
53. Shapiro B, Rambaut A, Drummond AJ. 2006. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol. Biol. Evol.* 23:7–9.
54. Shih SR, et al. 2000. Genetic analysis of enterovirus 71 isolated from fatal and non-fatal cases of hand, foot and mouth disease during an epidemic in Taiwan, 1998. *Virus Res.* 68:127–136.
55. Simmonds P. 2006. Recombination and selection in the evolution of picornaviruses and other mammalian positive-stranded RNA viruses. *J. Virol.* 80:11124–11140.
56. Simmonds P, Welch J. 2006. Frequency and dynamics of recombination within different species of human enteroviruses. *J. Virol.* 80:483–493.
57. Singh S, Poh CL, Chow VT. 2002. Complete sequence analyses of enterovirus 71 strains from fatal and non-fatal cases of the hand, foot and mouth disease outbreak in Singapore (2000). *Microbiol. Immunol.* 46: 801–808.
58. Solomon T, et al. 2010. Virology, epidemiology, pathogenesis, and control of enterovirus 71. *Lancet Infect. Dis.* 10:778–790.
59. Stamatakis A. 2006. RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.
60. Stanway G, et al. 2005. Family *Picornaviridae*, p. 757–778. In Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA (ed.), *Virus taxonomy*. Eighth report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, London, United Kingdom.
61. Tee KK, et al. 2010. Evolutionary genetics of human enterovirus 71: origin, population dynamics, natural selection, and seasonal periodicity of the VP1 gene. *J. Virol.* 84:3339–3350.
62. Tu PV, et al. 2007. Epidemiologic and virologic investigation of hand, foot, and mouth disease, southern Vietnam, 2005. *Emerg. Infect. Dis.* 13:1733–1741.
63. van der Sanden S, van der Avoort H, Lemey P, Uslu G, Koopmans M. 2010. Evolutionary trajectory of the VP1 gene of human enterovirus 71 genogroup B and C viruses. *J. Gen. Virol.* 91:1949–1958.
64. van der Sanden S, et al. 2011. Detection of recombination breakpoints in the genomes of human enterovirus 71 strains isolated in the Netherlands in epidemic and non-epidemic years, 1963–2010. *Infect. Genet. Evol.* 11: 886–894.
65. Wong SSY, Yip CCY, Lau SKP, Yuen KY. 2010. Human enterovirus 71 and hand, foot and mouth disease. *Epidemiol. Infect.* 138:1071–1089.
66. Yoke-Fun C, Abubakar S. 2006. Phylogenetic evidence for inter-typic recombination in the emergence of human enterovirus 71 subgenotypes. *BMC Microbiol.* 6:74.
67. Zoll J, Galama JM, van Kuppeveld FJ. 2009. Identification of potential recombination breakpoints in human parechoviruses. *J. Virol.* 83:3379–3383.